

音韻的作動記憶と英語熟達度の関係の検討

—— 日本人大学生を対象として ——

山 口 陽 弘 ・ 清 水 真 紀

群馬大学教育実践研究 別刷
第27号 281～290頁 2010

群馬大学教育学部 附属学校教育臨床総合センター

音韻的作動記憶と英語熟達度の関係の検討

—— 日本人大学生を対象として ——

山 口 陽 弘¹⁾・清 水 真 紀²⁾

1) 群馬大学教育学研究科専門職学位課程教職リーダー専攻

2) 高崎健康福祉大学

The Relationship Between Phonological Working Memory and English Proficiency of Japanese University EFL Learners

Akihiro YAMAGUCHI¹⁾ and Maki SHIMIZU²⁾

1) Program for Leadership in Education, Professional Degree Course,

Graduate School of Education, Gunma University

2) Takasaki University of Health and Welfare

キーワード：音韻的作動記憶、英語熟達度

Keywords: Phonological Working Memory, English Language Proficiency

(2009年10月30日受理)

1. はじめに

心理学, 言語学, 認知科学, 脳科学上の概念として, 人間の知的能力の中心的な位置を占めると言われている working memory というきわめて重要な構成概念がある。この語の日本語訳は「作業記憶」, あるいは「作動記憶」とその訳語が統一されていないが, 本稿では以下, 「作動記憶」と訳することにする。この作動記憶は, 短期記憶の概念を発展させたものであり, 厳密には両者は異なるが, ほぼ短期記憶と同種のものとみなしてよいものである。

短期記憶とは, まさに短期的にしか貯蔵されない記憶であり, たとえばランダムな数字の系列を読み上げられたとき, それを順序通りに再生できる場合の記憶のことを指す。

この短期的な記憶範囲検査によって測定される記憶範囲は, 成人の場合, 個人差はあるが 7 ± 2 程度であると言われている (Miller, 1956)。ミラーはこの数字

をマジカルナンバー 7 と呼び, 短期的に一度に処理できる最大の記憶可能な情報量であると, 半世紀以上前に論じていることは有名である。

この作動記憶が, 短期記憶と何が異なるかと言えば, 「短期記憶が情報の貯蔵機能を重視するに対し, 作動記憶は, 会話, 読書, 計算, 推理などの種々の認知機能の遂行中に情報がいかに操作され変換されるのかといった情報の処理機能を重視する」(森, 1999) ためにこの語を用いられることが多い。

作動記憶に関する仮説モデルも多く存在するが, 言語的情報処理のための音韻(「音声」とも訳されることもあるが, 本稿では以下「音韻」と訳す) ループ (phonological loop) と, 視覚的・空間的情報処理のための視空間スケッチ帳, この二つの下位システムを制御する中央制御部から構成されているというものが有名である (Baddeley, 1986)。なお, こうした作動記憶の仮説モデルに関しては実に多くの研究が存在するが, そのモデルを検討することは本稿では避ける。詳

細は三宅・齋藤（2001）の議論などを参照されたい。

本研究で特に問題にしたいのが、この言語的処理のために重要な音韻ループの働きに関わる「音韻的作動記憶（phonological working memory）」である。

この音韻的作動記憶は、母語の言語処理能力、獲得過程に限定されず、第二言語の処理過程においても大きく関わることが、上述の Baddeley を含めて多くの研究者が指摘している（e.g., Gathercole & Baddeley, 1993; de Jong, Seveke, & van Veen, 2000）。

しかしながら、日本人において外国語として最も広く受け入れられ、またその学習が学校教育の中でも必修として位置づけられることの多い英語学習における熟達過程と、この音韻的作業記憶との関係を直接的に検証した日本での実証的な研究は、それほど多くはないようである。

本研究では将来的には幼児の第二言語習得、イマージョン教育への適性判断のための判断材料にしていく可能性などを視野に入れている。しかし、今回はパイロットスタディとして、大学生である成人学習者の英語熟達度を測定し、それと音韻的作動記憶との関係を探索的に検討していくことを目的とする。

2. 音韻的作動記憶と英語熟達度との関係

音韻的作動記憶に関しても多くの議論があるが、本稿では、同種の研究を行っている湯澤（2008）と同様に、すでに述べたような Baddeley & Hitch（1974）の提唱した作動記憶モデルの下位システムである音韻ループの働きを指すものとする。この音韻ループ内において、記憶痕跡が時間の経過とともに減衰していくが、内的・外的に繰り返し構音化されることにより、保持されると想定されている。

音韻的作動記憶を実際に測定するためには、様々な方法が考えられる。上記で述べた数字の復唱課題などは古典的で分かりやすいものであるが、言語的な処理過程を測定するためには、非単語反復課題（nonword repetition task あるいは pseudoword repetition task とも呼ばれる）が、より純粋な音韻的作動記憶を測定するためには適切ではないかと考えられてきている。

この課題においては、課題遂行者は、音声提示された非単語（無意味綴りであるが、音韻構造的には当該言語でありうるようなもの）をできるだけ速く、正確

に復唱することが求められる。

一般に手続きが簡便であり、しかも記憶課題としては信頼性が高いとされるため（実際には後述するような諸問題が存在する）、年少の児童にも実施可能である。

この音韻的作動記憶の測定結果をもとに、将来的な外国語学習への適性が予測できるという実証的な研究例が、次のように数多く存在する。

Service（1992）は、英語を外国語として学ぶフィンランド人小学生を対象に研究を行い、音韻的作動記憶とその音韻的作動記憶測定時から2年半後の英語の成績間に相関があった（ $r = .66, p < .001$ ）としている。技能別では、リーディングで $r = .74 (p < .001)$ 、リスニングで $r = .62 (p < .001)$ 、ライティングで $r = .58 (p < .001)$ で、内容理解に関する技能、産出に関する技能といった違いにもかかわらず、中程度ないしは高い相関を示すことが明らかにされた。さらに、Service & Kohonen（1995）では、音韻的作動記憶と英語語彙習得との関係が示された。

Dufva & Voeten（1999）もフィンランド人小学生を対象として Service の素材を修正・追試して、やはり音韻的作動記憶の有効性を示している。構造方程式モデルを用いて分析した結果、2年次に測定された音韻的作動記憶の結果は、3年次の最後に測定された英語熟達度（語彙、リスニング、リーディング）に、21の強さで影響を与えていることが示された。

Cheung（1996）は、英語を外国語として学ぶ香港の中学生を対象として、確かに音韻的作動記憶の成績が語彙学習に関連していたという。つまり、音韻的作動記憶のよい人はそうでない人よりも効率よく新たな語を獲得できていたという。しかし、この有意な関連は、音韻的作動記憶の下位者にのみ限られみられた。

ギリシャ人の小学生を対象とした研究である、Masoura & Gathercole（1999）、Masoura & Gathercole（2005）でも、音韻的作動記憶と英語語彙知識との間に有意な相関があったと報告されている。それぞれ、英語学習歴が3～4年の学習者で、 $r = .43 (p < .05)$ と $r = .49 (p < .01)$ と中程度の相関を示した。

3. 本研究の目的

以上のように音韻的作動記憶は、外国語学習の成功

との間にポジティブな関連がみられることが、かなり頑健に指摘されている。しかし、この音韻的作動記憶には以下に示す重要な問題があるように思われる。

第一には、音韻的作動記憶を測定するための素材の問題である。この素材は、当該言語の音韻構造に即したものである必要がある。しかしこの同種の音韻構造というものを具体化することは難しいことである。つまり非単語素材に関して、英語に限定しても決定的に確立されたものが現時点では存在しないようなのである。すなわち、どのような素材が信頼性、妥当性があるのかがよくわかっていないという問題がある。

第二に、音韻的作動記憶は、どの程度言語間（たとえば日本語と英語）、素材間（非単語素材が研究者によってかなり違う）で変動するものなのか、そもそもよくわかっていない。作動記憶という構成概念の測定自体が、測定手法に大きく影響を受けるため（三宅・齋藤, 2001）、領域固有性がかなり強い構成概念ではないかという指摘が歴史的にそもそも存在する。

筆者らが危惧するのは、ここで言う領域固有性が、当該言語内で留まっており、少なくとも当該言語内では正の転移が大きいのであればよいのだが、当該言語内においても、特定素材における領域固有性が強ければ、音韻的作動記憶という構成概念自体の有効性が危ぶまれるのである。

以上の点から次のような調査目的を立てて、探索的に分析することとしたい。

- ①音韻的作動記憶の素材作り（日本語、英語）を行い、その信頼性を検討する。
- ②英語熟達度と音韻的作動記憶との関係を探索的に分析する。

なお、②の英語熟達度としては、先行研究で特に音韻的作動記憶との関連が指摘されている語彙知識を取り上げるとともに、リスニング能力にも焦点を当てる。

この②の部分を検討することは、いわば音韻的作動記憶課題の併存的妥当性を検討することになる。

4. 方法

4.1 実験協力者

日本人大学生英語学習者32名（内訳は、群馬大学教育学部生15名、高崎健康福祉大学学部生17名）が本実験に参加した。ただし、うち3名（いずれも高崎健康

福祉大学学部生）はすべての実験に参加していないので、以後、適宜分析から除いてある。

なお、実験に参加することは自由であることを確認した上、実験終了後、1,000～2,000円程度のUSBメモリまたは図書カードを謝礼として渡した。

4.2 素材と実験手続き

音韻的作動記憶を測定するものとして3種類の課題を、また英語熟達度を測定するものとして語彙サイズテストとリスニングテストを用いた。以下がその詳細である。

A) 英語熟達度の測定

○英語語彙サイズ測定テスト

日本人英語学習者用に作成され、また英語教育学の研究分野でも広く使用されている英語語彙サイズ測定テスト（望月・相澤・投野, 1998; 望月・相澤・投野, 2003, pp.212-221）を用いて、学習者の語彙知識を測定した。もとのテストでは1,000語レベルから7,000語レベルまで7段階あったが（徐々に難易度が増していく構成になっている）、本研究では、実験協力者の学習レベルを勘案して、1,000～5,000語レベルまでが実施された。各レベルの単語は30問用意され、合計150問が実施された。小集団で、35分程度で実施された。

○英語リスニングテスト

英語リスニングテストには、TOEIC Bridge のリスニングセクション (ETS, 2007a, pp.66-77) を用いた。この TOEIC Bridge は、「TOEIC テストよりも『易しく』『日常的で身近な』『時間の短い』初級学習者向けのテストとして」(ETS, p.10), ETS (Educational Testing Service) —TOEIC を開発している機関—が制作したものである。計50問で、Part 1 (写真描写問題) 15問, Part 2 (応答問題) 20問, Part 3 (会話問題) 15問という内訳になっていた。しかし、分析では、Part 別に分けられることはなく、これら全体でリスニング能力を構成するものとしてまとめて分析された。なお、この英語リスニングテストには25分間を要した。

以上2つの英語熟達度テストは、いずれもマークシート型の検査であり、1問1点で採点された。したがって、想定されるレンジは、英語語彙サイズ測定テストが0～150、英語リスニングテストが0～50であった。

B) 音韻的作動記憶の測定

音韻的作動記憶を測定する3種類の課題のうち、1つは学習者の母語（日本語）で行なわれ、2つは外国語（英語）で行なわれた。実験協力者にとっては、いずれも初めて接する課題であると考えられたため、練習を行なった後にこれらは行われた。

○課題1（日本語版）

まず、1つ目の日本語版（以下、「課題1」と呼ぶ）であるが、各非単語は、三文字の無意味綴りで、五十音のうち清音をランダムに組み合わせることで、本論文の第一著者および第二著者により作成された。その際、文頭が「ん」ないしは「を」となるものを除き、『大辞林』（松村、1995）でその三文字単語がないことを確認し、さらに両著者ともに、いずれも連想価が低く、無意味綴りであると認めたもののみを採用した（表1）。

表1 課題1の三文字の日本語無意味綴りの例

にはえ、かてせ、もかね、ほとき、つれむ、こねい、ひつら、すけね、れきあ、てあふ、へてす、むみほ、こへに、ぬたよ、めちな、ぬほえ、られの、ふらま、そてせ、ちうふ、まそあ、そわひ、ねるく、わえふ、ろこよ、れのき、もらま、つふぬ、ほせね、れらと、

このようにして作成された非単語は、2語～8語の範囲内で組み合わせられ、二語課題から八語課題までそれぞれ2セット用意された。たとえば、二語課題では、「にはえ」、「かてせ」と2語が音声で連続して聞こえてくるので、実験協力者はこれらを筆記再生することが求められた。同様に、三語課題では、「もかね」、「ほとき」、「つれむ」と3語が連続して聞こえてくるので、これらを筆記再生することが求められるといった具合であった。この音声化は、本稿の第二著者（日本語母語話者女性）によって行なわれ、CDに録音された。

なお、実験当初は、作成したもの全て、つまり八語課題まで実施しようと試みたが、実際に実施してみると五語課題以上の再生は非常に難しく、実験協力者に負担をかけることが分かり、そのため数名実施した時点で、それ以後の実験協力者には実施しなかった。また、結果は、その課題がどの程度できたかをあらわすスパン得点で算出した。スパン得点の算出法は、二語課題で1つもできなければ1.0、1つできれば1.5、2つできれば2.0である。二語課題ができた上で、三語課

題で1つもできなければ2.0のまま、1つできれば2.5、2つできれば3.0、以下同様に四語課題まで実施して、スパン得点を算出した。したがって、想定した課題1のスパン得点のレンジは1.0～4.0となる。

○課題2（英語版）

2つ目の課題は、Service (1992) および Service & Kohonen (1995) で使用された英語非単語の素材を使用した（以下、「課題2」と呼ぶ）。これは既に説明したように音韻構造としては英語と同様に聞こえるが、実際には存在しない非単語である。

具体的な素材を表2と表3に示す。これらは英語母語話者（イギリス人男性）によって発音され、CDに録音された。その際、非単語と非単語の間は、実験協力者がその間に口頭再生できるようにと、3秒間のポーズが置かれて編集された。以下の表も示すように、Practice Listを除けば全部で20個の非単語が用意された。

表2 課題2のPractice List

pilgey
gropulacture
peekrology
pungiligher
lendate

表3 課題2のList AとList B

List A	List B
redencable	disrire
mindon	hentidience
fentron	landipation
malegetron	dreplet
geplore	renditraction
punger	dupran
crentorious	fondet
nedrain	boxflab
mergimate	gontipastude
notifocal	bendestery

実施手続きであるが、表2のPractice Listで練習した後、表3のList A, List Bと順にCDが再生され、実験協力者は復唱を行なった。これらは、ICレコーダーに全て録音された。

この復唱結果は、第一著者、第二著者（いずれも日本語母語話者、第二著者は大学英語教員である）との間で協議され、「0；全く復唱できていない」、「1；復唱がある程度できているが完全ではない」、「2；ほぼ

完全に復唱できている」という三段階で評定された。いずれも音節が複数個ある非単語であるため、そのいずれも復唱できていなければ「0」、そのいくつかが復唱できていれば「1」、ほぼ完全な場合は「2」という評定を行った。この評定は、ICレコーダーに録音したものを再生して、二人の評定者で独立して行なった。評定者間一致率は、82.3%であった（実験協力者32名×20項目=640項目の中で不一致が113項目。したがって、 $527/640 \times 100 = 82.3$ ）。なお、不一致については、再度、ICレコーダーを何度も再生するなどしながら、二人で完全に合意を得られるまで協議して決定した。

この課題2の想定するレンジは、0～40であった。

○課題3（英語版）

Cheung (1996) が使用した素材を課題2と同様に作成し、実施した。

この課題の構造は、Practice List は課題2で使用したものと全く同じで(表2)、他のリストもほぼ表3と同じであるが、このような非単語を2語で復唱してもらったり、3語、4語、5語で復唱してもらうという構造が異なっていた。もとのCheungのものは5語以上の課題も存在したが、実験者の側で検討した結果、最大で5語で十分ではないかと思われたため、5語まで作成して実施した。

しかし、やはり数名まで実施した時点で、この課題が困難であり、実験協力者に負担をかけることがわかったため、実験を中止した。したがって、課題3は以下の分析からは除外してある。

課題1と課題2の実施時間は30分程度であった。実施方法はいずれも個別に実験室で行った。

また、A) およびB) 両方の課題とも、休憩時間を適宜取り、実験協力者が疲れないよう十分配慮して複数日をかけて実験を行った。

5. 結果

5.1 基礎属性

各テスト（課題）の記述統計は、表4のとおりである。英語リスニングテストの正解率は71.8%で、本研究で対象とした学習者にはやや易しい問題であったと言える。しかし、天井効果も床面効果もみられず、その点では適当な問題であったと考えられる。英語語彙サイズ測定テストも、満点150点に対して、最小値が80

点（（望月，1998）の語彙サイズ算出法に従うと2,667語レベルに相当）、最大値が137点（同様に、4,567語レベルに相当）となっていて、天井効果、床面効果ともにみられなかった。

音韻的作動記憶を測定した課題1（日本語版）は、満点が4.0であるのに対して、最大値がわずか2.5となり、かなり難しい課題であったと言える。一方で、課題2（英語版）は、得点が17～32点と広く散らばっていた。この課題1と課題2の違いは、2つの手続きの違いおよび得点の算出方法の違い（スパン得点であったかそうでなかったか）を反映したものと思われる。

なお、いずれも分布は正規分布に近く、著しい歪みはなかった。

表4 各テスト（課題）結果の記述統計

テスト・課題（レンジ）	N	M	SD	Min	Max
英語リスニングテスト（0-50）	29	35.9	5.3	23	46
英語語彙サイズ測定テスト（0-150）	32	110.7	14.7	80	137
音韻的作動記憶 課題1（日本語版）（1.0-4.0）	32	1.6	0.4	1	2.5
音韻的作動記憶 課題2（英語版）（1-40）	32	26.3	3.8	17	32

5.2 基礎属性間の相互相関

基礎属性間でピアソンの積率相関分析を行なった。その結果、有意な相関があったものは、語彙サイズとリスニングとの関係のみ（ $N=29$, $r=.468$, $p<.05$ ）であった。

他は、リスニングと日本語音韻的作動記憶との間に負の有意傾向があった（ $N=29$, $r=-.313$, $p<.10$ ）が、ほかはいずれも±0.1以下の相関であり、まず無相関と言ってもよい結果となった。

5.3 音韻的作動記憶課題（英語版）の項目分析

（正答率・標準偏差・IT分析）

音韻的作動記憶を測定した課題2（英語版）について、各項目がこの課題の総合得点とどのような関係があるかを探るため正答率と標準偏差を算出した。その結果をまとめたものを表5に示す。この中で特に正答率が低い項目（平均正答率が1.0を切るもの）はmalegetron (0.63), dreplet (0.94), renditraction (0.44), boxflab (0.84) であった。

次に合計得点と項目間の相関を分析するIT分析を

表5 音韻的作動記憶課題2の非単語ごとの正答率と標準偏差(()内)の結果

redencable 1.66 (0.602)	mindon 1.44 (0.619)	fentron 1.00 (0.508)	malegetron 0.63 (0.554)
geplere 1.56 (0.504)	punger 1.13 (0.421)	crentorious 1.22 (0.659)	nedrain 1.53 (0.671)
mergimatate 1.50 (0.568)	notifocal 1.31 (0.592)	disrire 1.75 (0.440)	hentidience 1.31 (0.592)
landipation 1.50 (0.718)	dreplet 0.94 (0.435)	renditraction 0.44 (0.619)	dupran 1.75 (0.440)
fondet 1.84 (0.369)	boxflab 0.84 (0.369)	gontipastude 1.59 (0.560)	bendestery 1.38 (0.609)

表6 音韻的作動記憶課題2の非単語ごとのIT相関分析の結果

redencable 0.437	mindon 0.370	fentron 0.591	malegetron 0.282
geplere 0.061	punger 0.413	crentorious 0.327	nedrain -0.216
mergimatate 0.254	notifocal 0.000	disrire 0.109	hentidience -0.199
landipation 0.296	dreplet 0.177	renditraction -0.004	dupran 0.337
fondet 0.248	boxflab 0.223	gontipastude 0.348	bendestery 0.343

行なった。その結果をまとめたものが表6である。値が高いものほど、総合得点との関係が強いことを表している。ここで仮に基準を+.20とすると、それを超えるものは全体の20項目のうち13項目となった。他方、+.20に満たないもの(負の値であったものも含む)は7項目あった。

とリスニング能力)であり、これは中程度の相関があった。しかし、これは本研究の最低限度の妥当性を確認するための、当然な正の相関であり、実験協力者がランダムな反応はしていなかったということを示すに過ぎない消極的なものである。

どうしてこのようなことが起きたかを、以下検討していく。

6. 考察および再分析結果と今後の研究方針

6.1 課題・テスト間の相関関係

探索的な作業仮説として、英語熟達度、特にリスニング能力と音韻的作動記憶との間には相関があるのではないかと考えて実験を行ったが、今回の結果からは特に相関がみられなかった。むしろ日本語での音韻的作動記憶との間に負の相関(有意傾向)という結果もあり、これは予想と反する結果である。

また日本語での音韻的作動記憶と英語での音韻的作動記憶との間に中程度の正の相関があるかと思われたが、ほぼ無相関という結果であった。

唯一相関があったのが、英語熟達度間(語彙サイズ

6.2 各課題の信頼性の低さ

以上のように、基本的に想定していた各種の相関が出ない原因として、各課題の信頼性が低いのではないかと考え、その信頼性を各々計算したところ、リスニング課題の $\alpha = .698$ (50項目)、語彙サイズ課題の $\alpha = .917$ (150項目)、音韻的作動記憶課題1(日本語版)の $\alpha = -.072$ (6項目)、音韻的作動記憶課題2(英語版)の $\alpha = .600$ (20項目)であり、いずれもその項目数から考えると、心理学実験で使用する素材としては大変低い。本研究で出るべき基礎属性間の相関が出なかった最大の要因はここにあるようである。

6.3 音韻的作動記憶課題（日本語版）の素材の再検討の必要性

特に日本語での作動記憶に関しては、レンジが狭かったということ、項目数が少なかったということもあるが、通常の心理検査としては体をなしていないほど今回信頼性が低かった。素材作成の時点で多くの問題があったことが反省点としてあげられる。

第一に、問題の難易度が全般に高く、日本語での素材を、一語課題から提示する必要性も感じられた（二語課題でも完全にできない実験協力者が存在した）。第二に、二語課題も、今回は2セットの提示であったが、この数が信頼性を確保するためには足りなかったのではないかと考えられる。次回からは少なくとも4つ程度以上の課題を提示する必要性を感じた。

今回の実施に当たっては項目数を6つではなく、英語の音韻的作動記憶課題と同様に、20項目以上に増やして再実施する必要があるだろう。今回、予期せぬ負の相関（リスニング能力との間）が生まれた最大の原因は、日本語の音韻的作動記憶課題自体の信頼性の欠如であったと推定される。

また、場合によっては日本語の音韻的作動記憶も、筆記再生ではなく、口頭での復唱による再生にして検討すべきかもしれない。筆記する力を大学生であるため、そこまでの必要がないと想定していたが、それが問題であったのかもしれない。

さらに言えば、筆者らが日本語に関しての音韻構造に関して十分検討を行わずに素材作成を行ってしまったことも今回の失敗の原因であろう。表1で示したような素材の中でも、日本語の音韻構造としてありえない（あまりない）ものが入っている可能性も十分ありうる。たとえばエ音の後に続くイ音などは、日常用語としては長音で発音することが現代日本語では一般化している（たとえば「ケイタイ」を「ケータイ」と発音する）ようだが、それゆえ「こねい」などの正答率が低かったのかもしれない。もちろん、録音時には入力者はきちんと発音しているのだが、聞きとり側の体制として長音化して記述しやすくなっているのかもしれない。

再実験に当たっては、日本語音韻構造に関しても再検討を行い、素材作りの際に、日本語に即した素材を作成するよう努力するつもりである。

6.4 ハイスティクスなテストの信頼性の低さへの注意喚起

それ以外の音韻的作動記憶課題（英語版）、英語リスニングテスト、英語語彙サイズ測定テストに関しては、最低限度の信頼性は確保されていると言える。しかし、項目数、さらにはそれが能力検査である点などをかんがみると、決して高い値ではない。

一般に心理学的なテストの場合は、0.7以上の信頼性は必要とされ、ハイスティクス (high-stakes) なテスト（当該のテスト結果が受検者の将来に大きく影響を与えるテスト）である場合ならば、0.9以上は欲しい。

すなわち予測力、再現性という点で考えると0.7であればその二乗が0.49であり、ほぼ半分程度の予測力（再現性）しか備わっていないことになる。その程度の信頼性で人生を決定することは危険ではないかとされるのである。

特に英語リスニングテストの信頼性の低さ（ $\alpha = .698$ ）は、これが TOEIC Bridge の開発機関でもある ETS から公にされているものであり、またハイスティクステストであると考えられる TOEIC の初級版としても位置づけられているだけに気になる点である。なお、*TOEIC Bridge Examinee Handbook* (ETS, 2007b, p.18) によれば、リスニングセクション、リーディングセクションそれぞれの信頼性は KR-20 によれば「おおそ0.85以上」(approximately 0.85 and up) と公表されている。今回の結果がたまたま実験協力者の等質性（リスニング能力に個人差がなかった可能性もある）に基づいての信頼性の低さであるのか、あるいはそもそも当リスニングテスト自体がそれほど信頼性が低いものなのかは、再考すべき点ではあるが、TOEIC Bridge のリスニングテストの信頼性の低さに関しては、もう少し注意喚起しておきたいし、今後の研究が必要な点であろう。

6.5 音韻的作動記憶課題（英語版）の素材の検討

英語素材での音韻的作動記憶の測定が、今回の研究の主たる目的でもあるため、今回の素材を用いて、 α 係数を上げるため、IT 分析などを行った。その際、全体相関の低いものを削除して項目精選を行い、さらに相関分析を探索的に行なった。

現時点では、英語の音韻的作動記憶の素材自体が十分標準化されているとは言えない段階であるため、こ

れを精選していくためにも、全体-部分相関の低いものを検討しておくことは有益であろう。

以下の表 7 に、あらためて全体との相関が 0.2 未満であった素材を示す。括弧内は全体得点と、当該素材との相関を示している。すなわち、これらの素材は、非単語でも適切か否かということが、「日本人大学生にとっては」であるが、要注意の素材ということである。このうち renditraction と dreplet とは正答率も 1.0 未満と低く、そもそも項目として「日本人・成人大学生」にとっては妥当ではない可能性が高い。

表 7 IT 相関係数の低かった非単語素材

notifocal (.000)
renditraction (-.004)*
geplere (.061)
disrire (.109)
dreplet (.177)*
hentidience (-.199)
nedrain (-.216)

註) *は正答率も 1.0 未満で低かったもの

なお、各素材を、音節数という点から検討したところ、notifocal, renditraction, hentidience の 3 つは 4 音節で構成されていて(音節数が多い)、その他 4 つの非単語は 2 音節で構成されていた(音節数が少ない)。この Service(1992)および Service & Kohonen(1995)の非単語素材が、2 音節ないしは 4 音節のいずれかであったことを考慮すると、必ずしも、この表 6 に掲げられた素材が、音節数が多いものに偏っていた、あるいは音節数が少ないものに偏っていたと言うことはできないことが分かる。

ともあれ、上記の項目を削除することで、今回の音韻的作動記憶課題(英語版)の α 係数を 0.748 (13 項目)に上昇させることはできる。しかし、そのような方法で修正しても、英語リスニング能力、英語語彙サイズとの間に相関はみられず、項目精選の効果は全くみられなかった。この結果、素材を精選する方針を避けて、むしろ項目数が多くなることで実験協力者への負担は増えるが、現時点ではもう少し項目数を増やして幅広く素材を確保していく方針で、信頼性を上昇させる必要があるようである。

6.6 CNRep との関係

そもそも英語の音韻的作動記憶課題であるが, Gathercole & Baddeley (1996) により非単語反復課題で、

一応標準化されたものが既に作成されている。これは、英語母語話者向けに作成されたものであり、非単語反復課題である点など本研究で用いた課題と根本的部分は類似しているが、刺激の構造が少し異なる。Gathercole & Baddeley による上記テストは CNRep と通称では言われるので、以下はそう呼ぶ (the Children's test of Nonword Repetition の頭文字を取って命名されている)。

刺激は 2 音節, 3 音節, 4 音節, 5 音節からなる英語の非単語であり、各 10 個, 計 40 個からなる。このように音節数にバリエーションがある点と、項目数が多い点が本実験の素材と大きく異なる点である。

これらの非単語は一つずつランダムに聴覚提示され、実験協力者は提示された非単語をそのまま声に出して反復するように求められる点は同じである。

CNRep の技術的な情報をいくつかまとめておく。これはイギリス人の 612 名の 4 歳から 8 歳までの児童を対象に標準化されている。CNRep の再検査信頼性は 0.77, 折半法での信頼性は 0.66 という報告がなされている (Gathercole & Baddeley, 1996)。項目数が多い割には信頼性があまり高くない点が気になることであるが、児童を対象としている点からはやむを得ないのかもしれない。

もう一点、本研究と刺激素材としての大きな違いは、英語非単語を音節ごとに児童の音声反応が正しいか否かを判定する点である。本研究では、非単語全体を 3 段階で評価しており、音節にも注目しているが、音節単位での評価ではなかった。

そもそも CNRep が英語圏での英語非単語の音韻的作動記憶を測定しているものであり、他言語圏、さらには児童以外での利用可能性までを保証して作成されているものではない。

日本でも湯澤・湯澤・関口 (2009) によって CNRep をインターナショナルスクールに通う日本人幼児 (4 歳~6 歳) 25 名を対象に実施しており、他言語圏での CNRep の利用可能性を示しているが、その結果に関しては、明確なものとは言い難いようである。というのは、CNRep の妥当性を検証するために音韻認識課題を湯澤らは別途同時に実施して、併存的妥当性を検証しようとしているが、残念ながら明確な正の相互相関が出ていないのである。

すなわち CNRep の信頼性, 妥当性の検討もいまだ日本での児童, 成人に対しては研究途上にあると言える。次回の実験に当たっては CNRep も並行して実施する予定である。

6.7 今後の課題

本研究の音韻的作動記憶課題 (英語版, 日本語版両方) の信頼性, 妥当性を, 今後さらに上げるためには, まず CNRep と今回用いた Service (1992) および Service & Kohonen (1995) で使用されているもの (さらには今回は実験途中で中止した Cheung (1996) もあわせて) との併存的妥当性を検証していくことが重要であろう。ただし, 日本語母語話者 (Service の課題はフィンランド語母語話者ではその有効性は示されているが) にもこれらの音韻的作動記憶課題がそもそも適用可能かどうかはまだ未確定と言うべきであるため, 並行して両方の素材の探索的研究をしていくべきである。

今回の実験方法で反省すべき点として, CNRep の評定方法・得点化の方法との違いもある。CNRep が音節単位での評定であるということもあるが, 日本での実施に当たっては, 湯澤ら (2009) の研究においては, 英語母語話者 (イギリス人) がこの評定を行っている。この評定者はイギリス国内大学の心理学部において音声学・言語心理学のコースを終了しており, 発音記号の読み方や音素・音節についての音声の聞きとりに関して熟知していたという。

本研究では第一著者は英語教育学に関する論文も書いており, 発音記号の読み方, 音素, 音節に関しての基礎知識はあるが, 英語教育に関しては専門ではなく, 英語母語話者ではないどころか, 教育心理学が専門である。

第二著者も英語教育の専門家ではあるが英語の母語話者ではない。2 人の合議による評定自体に問題があった可能性は十分ある。英語母語話者に課題の評定をしてもらうことも必要かもしれない。

また, 振り返って根源的な問題を述べると, 「音韻構造」として当該言語と同じであるということを言語学的に定義することは, 根源的に非常に難しい問題をはらんでいる。

既に実在する単語の一部をうまく改変して非単語を作成したとしても, もとの単語との連想価は必然的に

生じる。そもそも何らかの実際の単語との連想価が全く存在しないような非単語というのは原理的に存在しないであろう。

既に述べたように, 今回筆者ら自らが作成した日本語素材にしても, 連想価は低いと想定したものではあるが, それにしても相対的に連想価が高いものも存在し, その高いことが正確な復唱に対して, 正に作用することもあれば負に作用することも想定される。筆者らが日本人母語話者であったから非単語をうまく作成できるとは限らないのである。

いわんや大学生においては, 英語の非単語で熟知価が高い語とそうでない語とが, 剰余変数として大きな要因になる。この際に, 個人の既有知識との関係が正に働くのか, 負に働くのかが現時点では不明である。これが中学から高校にかけての英語教育の影響によって, 大きな剰余変数として働いた可能性が否めない。

CNRep や今回使用した Service 版, Cheung 版の非単語も, 正確な音韻構造に関する知識をもとにして作成されたか否かは原論文を検討しても少し怪しいようである。

今回の不安定な結果も, あえて筆者らの考察を述べれば, 大学生はむしろ既有知識としての英単語 (さらに言えば日本語の単語も同様) があるため, そこに引きずられて正確な復唱が難しかった可能性が考えられる。すなわち, 幼児・児童よりも非単語課題を復唱することが難しい場合も十分生じる。しかもその阻害要因が個人の既有知識によって大きく異なるため, 単純に非単語復唱という方法では音韻的作動記憶を測定することが難しいのかもしれない。

6.8 まとめ

今後の研究方針, 再実験に当たっての方針を最後にまとめておく。①日本語音韻的作動記憶課題の充実 (項目数, 単語の精選, 復唱方法の検討, 音韻構造の再検討), ②英語音韻的作動記憶課題の充実 (語数, 単語の精選, CNRep との併存的妥当性の検討, 評定方法の再検討, 音韻構造の再検討), ③英語熟達度の測定課題の検討 (英語リスニングテスト, 英語語彙サイズ測定テストをより充実させる), ④実験協力者の再検討 (大学生から幼児・児童へ) といった点が, 今後の課題としてあげられる。

何よりも当該言語に関する音韻構造に関して自然か

否かという視点が今回の研究には欠けていたように思う。この点を再検討, 吟味した上で再実験を行う予定である。

引用文献

- Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp.47-90). New York: Academic Press.
- Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology*, 32, 867-873.
- de Jong, P. F., Seveke, M. J., & van Veen, M. (2000). Phonological sensitivity and the acquisition of new words in children. *Journal of Experimental Child Psychology*, 76, 275-301.
- Dufva, M., & Voeten, M. J. M. (1999). Native language literacy and phonological memory as prerequisites for learning English as a foreign language. *Applied Psycholinguistics*, 20, 329-348.
- ETS (2007 a). 『TOEIC 公式ガイド&問題集』. 国際ビジネスコミュニケーション協会・TOEIC 運営委員会.
- ETS. (2007b). *TOEIC Bridge examinee handbook*. Retrieved October 28, 2009, from <http://www.ets.org/Media/TOEIC%20BridgeExam.pdf>
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. East Sussex, UK: Psychology Press.
- Gathercole, S. E., & Baddeley, A. D. (1996). *The children's test of nonword repetition*. London: Psychological Corporation.
- Masoura, E. V., & Gathercole, S. E. (1999). Phonological short-term memory and foreign language learning. *International Journal of Psychology*, 34, 383-388.
- Masoura, E. V., & Gathercole, S. E. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory*, 13, 422-429.
- 松村 明 (編) (1995). 『大辞林』 (第二版). 三省堂.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- 三宅 晶・齋藤 智 (2001). 「作動記憶研究の現状と展開」. 『心理学研究』, 72, 336-350.
- 望月正道 (1998). 「日本人学習者のための英語語彙サイズテスト」. 『語学教育研究所紀要』, 12, 27-53.
- 望月正道・相澤一美・投野由紀夫 (2003). 『英語語彙の指導マニュアル (英語教育21世紀叢書)』. 東京: 大修館書店.
- 森 敏昭 (1999). 「作動記憶」. 中島義明ほか (編). 『心理学辞典』 (p.299). 有斐閣.
- Service, E. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology*, 45, 21-50.
- Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, 16, 155-172.
- 湯澤美紀. (2008). 「日本人幼児における英語の音韻認識の形成に関する発達的变化」. 『発達研究』, 22, 173-180.
- 湯澤美紀・湯澤正通・関口道彦 (2009). 「日本人幼児の英語音韻習得のプロセスに関する研究: 音韻的作動記憶, 音韻認識, 日本語語彙量からの検討」. 『発達研究』, 23, 189-200.
- (本研究は, 第一著者に対して交付された, 『平成21年度ぐんま国際教育財団の語学助成金「ぐんま国際アカデミー」における入学適性のための選抜法の改善案の検討II』の一部を使って行われた.)

(やまぐち あきひろ・しみず まき)